F & F

Access DB# 126758

SEARCH REQUEST FORM

29

Scientific and Technical Information Center

Requester's Full Name: Jacob F. Bellinger Examiner #: 80128 Date: 9 July 2004

Art Unit: 2175 Phone Number 305 3775 Serial Number: 10/054,366

Mail Box and Bldg/Room Location: 4A30 Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

****

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: System & Method for detecting duplicate and similar documents

Inventors (please provide full names): James' Cooper, Anni Coden, Eric Brown

Earliest Priority Filing Date: 1/22/2002

For Sequence Searches Only *Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.*

sorting documents by size then comparing only documents that are ~~around~~ around the same size (no more than a predetermined size larger)

(claims 8 and 9)

****

STAFF USE ONLY	Type of Search	Vendors and cost where applicable
Searcher: Terence Esterfield	NA Sequence (#)	STN
Searcher Phone #: 308-7759	AA Sequence (#)	Dialog
Searcher Location: 4B30	Structure (#)	Questel/Orbit
Date Searcher Picked Up: 7/9/04 11:15 am	Bibliographic	Dr.Link
Date Completed: 7/9/04 3:30 pm	Litigation	Lexis/Nexis
Searcher Prep & Review Time:	Fulltext	Sequence Systems
Clerical Prep Time:	Patent Family	WWW/Internet
Online Time:	Other	Other (specify)

PTO-1590 (8-01)

Set	Items	Description
S1	1054100	DOCUMENT OR RECORD? ? OR FILE? OR TEXT?
S2	1529161	SORT? OR ORDER? OR RANK? OR ARRANG? OR ORGANIZ? OR RELEVANCE? OR POSITION?
S3	1429620	SIZE? OR SIZING OR THRESHOLD OR LENGTH OR CAPACITY? OR LIMIT? OR EXTENT
S4	845227	COMPAR? OR MATCH? OR EQUATE?
S5	1517765	SIMILAR OR ALIKE OR COMPARABLE OR CORRESPOND? OR EQUIVALENT OR SAME OR EQUAL OR IDENTICAL OR PREDETERMIN? OR AROUND
S6	1270408	RANGE OR EXTENT OR REALM OR SCOPE OR SPHERE OR AMOUNT OR TOTAL OR TOTALITY OR SMALLER OR LARGER
S7	79190	S1 (S) S2 (S) S3
S8	2457	S4 (S) S1 (S) (S5 (3N) S6)
S9	530	S7 (S) S8
S10	122	S9 AND IC=G06F?
S11	3	S10 AND IC=G06F-007?
S12	119	S10 NOT S11

File 348:EUROPEAN PATENTS 1978-2004/Jul W01
(c) 2004 European Patent Office
File 349:PCT FULLTEXT 1979-2002/UB=20040701,UT=20040624
(c) 2004 WIPO/Univentio

11/5,K/1 (Item 1 from file: 349)
DIALOG(R)File 349:PCT FULLTEXT

00857187 **Image available**

METHOD AND SYSTEM FOR ORGANIZING OBJECTS ACCORDING TO INFORMATION CATEGORIES

METHODE ET SYSTEME D'ORGANISATION D'OBJETS EN FONCTION DE CATEGORIES D'INFORMATION

Patent Applicant/Assignee:
TZUNAMI INC, c/o Aaron Etra, Martin & Taub LLP, 1350 Avenue of the Americas, New York, NY 10019, US, US (Residence), US (Nationality), (For all designated states except: US)

Patent Applicant/Inventor:
KLEINBERGER Paul, 4 Hamaapilim Steet, 92545 Jerusalem, IL, IL (Residence), IL (Nationality), (Designated only for: US)
BEZEM Shlomo, 697 Har Ramon Street, 71908 Macabim, IL, IL (Residence), IL (Nationality), (Designated only for: US)
EZER Yoav, 27 Alkihi Street, 93807 Jerusalem, IL, IL (Residence), IL (Nationality), (Designated only for: US)
HACHLILI Guy, 32 Alkihi Street, 93807 Jerusalem, IL, IL (Residence), IL (Nationality), (Designated only for: US)
JACOBSON Ron, 1 Brazil Street, 69710 Tel Aviv, IL, IL (Residence), IL (Nationality), (Designated only for: US)
SANDLERMAN Nimrod, 44 Churgin Street, 52356 Ramat Gan, IL, IL (Residence), IL (Nationality), (Designated only for: US)
ZILBERMAN Doron, 4 Rabinowitz Street, 58672 Holon, IL, IL (Residence), IL (Nationality), (Designated only for: US)

Legal Representative:
COLB Sanford T (et al) (agent), Sanford T. Colb & Co., P.O. Box 2273, 76122 Rehovot, IL,

Patent and Priority Information (Country, Number, Date):
Patent: WO 200190840 A2-A3 20011129 (WO 0190840)
Application: WO 2001IL486 20010525 (PCT/WO IL0100486)
Priority Application: US 2000208110 20000526

Designated States: AE AG AL AM AT AU AZ BA BB BG BR BY BZ CA CH CN CO CR CU CZ DE DK DM DZ EC EE ES FI GB GD GE GH GM HR HU ID IL IN IS JP KE KG KP KR KZ LC LK LR LS LT LU LV MA MD MG MK MN MW MX MZ NO NZ PL PT RO RU SD SE SG SI SK SL TJ TM TR TT TZ UA UG US UZ VN YU ZA ZW
(EP) AT BE CH CY DE DK ES FI FR GB GR IE IT LU MC NL PT SE TR
(OA) BF BJ CF CG CI CM GA GN GW ML MR NE SN TD TG
(AP) GH GM KE LS MW MZ SD SL SZ TZ UG ZW
(EA) AM AZ BY KG KZ MD RU TJ TM

Main International Patent Class: **G06F-017/00**
International Patent Class: **G06F-007/00**
Publication Language: English
Filing Language: English
Fulltext Availability:
Detailed Description
Claims
Fulltext Word Count: 35125

English Abstract
A method and system of organizing items including building up clusters of items, each item having information associated therewith, during building up of the clusters evaluating dynamically a metric of the cluster, the metric of the cluster expressing at least whether the items in a cluster have more in common with each other than they have in common with items outside of the cluster.

French Abstract
L'invention concerne une methode et un systeme permettant d'organiser des articles. Selon cette methode et ce procede, des groupes d'articles sont crees, chaque article presentant une information qui lui est propre. Au cours de la creation des groupes, cette methode et ce systeme evaluent un parametre du groupe de maniere dynamique, ce parametre du groupe

déterminant au moins si les articles d'un meme groupe ont plus en commun les uns avec les autres, qu'avec les articles n'appartenant pas audit groupe.

Legal Status (Type, Date, Text)
Publication 20011129 A2 Without international search report and to be republished upon receipt of that report.
Search Rpt 20020228 Late publication of international search report
Republication 20020228 A3 With international search report.
Examination 20030109 Request for preliminary examination prior to end of 19th month from priority date

Main International Patent Class: **G06F-017/00**
International Patent Class: **G06F-007/00**
Fulltext Availability:
Claims

Claim
... create several
D
alternative groups of clusters and select said best group.
0
100. A method of **organizing** information comprising:
breaking down clusters of information items;
during breaking down of the clusters evaluating dynamically ...each other than they have in cornmon with items outside of the cluster. 101. A method of **organizing** inforrnation according to claim 1 00 and wherein said metric is a commonality metric. 102. A method of **organizing** information according to claim 100 and wherein said metric is a similarity metric. 1 0 31. A method of **organizing** information according to claim 100 and wherein said metric is a non-cornmonality metric. 104. A method of **organizing** information aceording to claim 100 and wherein said metric is a non-similarity metric. 105. A method of **organizing** infonnation according to claim 100 and wherein each item includes at least one descriptor and said metric...a most preferred cluster has the highest Cluster Quality Metric of all possible first clusters available for **comparison** . 149. A method according claims 139 - 147 and wherein a structure of clusters is presented to the...number of iterations haye taken place.
1 5
164. A method according to claim 161 and wherein **limitation** of calculations to qualified descriptors are used for calculating a Cluster Quality Metric CQM:
CQM = aX + bY...

...All unique descriptors of the items of the collection are identified;
Step (b). The identified descriptors are **ranked** aecording to their popularity in the collection;
Step (c). A "base item" is chosen as a first item of a "base: cluster";
Step (d). A plurality of " **comparison** item? are chosen;
Step (e). The base item is considered to be a first item in a "base cluster", and each **comparison** item is considered to be a first item in a " **comparison** cluster";
Step (f). The base cluster, now including all items of the collection having a higher gravity score with respect to the base cluster than with respect to any of 5 the **comparison** clusters, is retained as the desired preferred cluster for the collection.
166. A method according to claim 161 and wherein. said identified unique descriptors are the highly **ranking** descriptors and wherein descriptors that exist in many items of the collection of items are **ranked** above descriptors existing in few items of the collection. 167. A method according to claim 166 and wherein each descriptor receives a **rank** score equal. to the number of items of the collection in which said descriptor exists. 168. A method according to claim 167 and wherein said **ranking** is influenced by a weighting factor dependent on some characteristics of the descriptors 169. A method according to claim 168

and wherein. said **ranking** is influenced by a weighting factor dependent on some characteristics of the items in which they appear. 170. A method according to claim 169 and wherein said **ranking** is influenced
55
by a weighting factor dependent On Some characteristics of descriptors of items having few...

...than descriptors of items having many descriptors. 171. A method according to claim 170 and wherein said **ranking** is influenced by a weighting factor dependent on some characteristics of descriptors which are nouns are given having the highest- **ranking** combination of high-**ranking** descriptors.
173. A method according to claim. 172 and wherein said **ranking** is accomplished by first calculating an item. score for each item, which is the sum of the...

...the highest item. score.
175. A method according any- of claims 100 - 174 and wherein. a first **comparison** item is an item having a high item score, yet also having a low similarity
score when **compared** to the base itern; and
additional **comparison** items are chosen, being items having a high item score, yet also having a low similarity score when. **compared** to the base item and further having a low similarity score when **compared** to all previously chosen **comparison** items.
176. A method according to claim. 165 the method comprising selecting a base cluster and a plurality of **comparison** clusters, each of these clusters having a single item, wherein in step (e) a gravity score is...

...for each item of the collection with respect to said base cluster and with respect to each **comparison** cluster, and each item is added to said cluster with respect to which. it has the highest...

...claims 100 - 178 and wherein a directory
tree is created automatically for the results of a free **text** search
180. A method according to claim 179 and wherein said directory tree enables the user to...

...based on several subjects.
186. A method according to claim 179 and wherein said directory tree is **organized** , and the information items are **sorted** into the directory tree, based on commonality metric that involves several terms.
187. A method according to claim 179 and wherein said directory tree is **organized** , and the information (inverted exclamation mark)teras are **sorted** into the directory tree, based on a commonality metric that involves tenns that were not specified by...

...the user's query.
188. A method aceording to claim 179 and wherein said directory tree is **organized** , and the information items are **sorted** into the directory tree, based on a 1 5 commonality metric that involves a plurality of terms.
189. A method according to claim 179 and wherein said directory tree is **organized** , and the information items are **sorted** into the directory tree, based on a metric of lack of commonality between information items.
190. A method aceording to claim 179 and wherein said directory tree is **organized** , and the information items are **sorted** into the directory tree, in an iterative manner where information items are added or removed from clusters...

...of clusters, create several alternative groups of clusters and select said best group.
199. A method of **organizing** inforination comprising:
changing the population of clusters of information items,
during changing the population of the clusters...each other than they have in conunon with items outside of the cluster. 200. A method of

organizing information according to claim 199 and wherein said metric is a cornrnonality metric.
59
. . A method of **organizing** information according to claim 199 and wherein said metric is a similarity metric. 202. A method of **organizing** information according to claim 199 and wherein said metric is a non-commonality metric. 203. A method of **organizing** information according to claim 199 and wherein said metric is a non-similarity metric. 204. A method of **organizing** infonnation according to claim 199 and wherein each item includes at least one descriptor and said metrie...

...of claims 205 - 210 and wherein calculating said similarity score includes assigning at least one of a **match** count and an unrnatch count to a pair of items.
0
212. A method according to claim, 211 and also comprising weighting at least one of said **match** count and said unmatch count.
213. A method according to claim 211 and wherein said metric includes a 5 metric which is equal to the weighted **match** count.
214. A method according to claim 211 and wherein. said metric includes a metric which is...205 - 229 and wherein
an intra cluster gravity score ICGS is calculated and wherein said ICGS is **equal** to the **total** of the gravity scores for each item in said cluster with respect to all items inside
z...

...a most
preferred cluster has the highest Cluster Quality Metric of all possible first clusters available for **comparison** . 248. A method according claims 238 - 246 and wherein a structure of clusters is presented to the...a set number of iterations have taken place. 263. A method according to claim 260 and wherein **limitation** of calculations to qualified descriptors are used for calculating a Cluster Quality Metric CQM:
CQM = aX + bY...

...All unique descriptors of the items of the collection are identified;
Step (b). The identified descriptors are **ranked** according to their popularity in the collection;
Step (c). A "base itern" is chosen as a first item of a "base cluster";
Step (d). A plurality of " **comparison** iteras" are chosen;
Step (e). The base item is considered to be a first item in a "base cluster", and cach **comparison** item is considered to be a first item in a " **comparison**
cluster";
Step ffi. The base cluster, now including all items of the collection having a higher gravity score with respect to the base cluster than with respect to any of the **comparison** clusters, is retained as the desired preferred cluster for the collection.
265. A method according to claim 260 and wherein said identified unique descriptors are the highly **ranking** descriptors and wherein descriptors that exist in many items of the collection of items are **ranked** above descriptors existing in few (inverted exclamation mark)tems of the collection.
69
. A method according to claim 265 and wherein each descriptor receives a **rank** score equal to the number of items of the collection in which said descriptor exists. 267. A method according to claim 265 and wherein, the **ranking** is influenced by a weighting factor dependent on some characteristics of the descriptors 268. A method according to claim 266 and wherein said **ranking** is influenced by a weighting factor dependent on some characteristics of the items in which they 0 appear. 269. A method according to claim 268 and wherein said **ranking** is influenced by a weighting factor dependent on some characteristics of descriptors of items having few descriptors...
...than descriptors of items having many descriptors. 270. A method

according to claim 269 and wherein said **ranking** is influenced by a weighting factor dependent on some characteristics of descriptors which are nouns are given item having the highest- **ranking** combination of high- **ranking** descriptors.
272. A method according to claim 271 and wherein said **ranking** is accomplished by first calculating an item score for each item, which is the sum of the...

...highest item score.
274. A method according any of claims 199 - 273 and wherein a first
70
comparison item is an item having a high item score, yet also having a low similarity
score when **compared** to the base item; and
additional **comparison** items are chosen, being items having a high item score, yet also having a low similarity score when **compared** to the base item and further having a low similarity score when **compared** to all previously chosen **comparison** items.
275. A method according to claim 264 the method comprising:
selecting a base cluster and a plurality of **comparison** clusters, each of these clusters having a single item, wherein in step (e) a gravity score is...

...for each item of the collection with respect to said base cluster and with respect to each **comparison** cluster, and each item is added to said cluster with respect to which (inverted exclamation mark)t...

...claims 219 - 277 and wherein a directory
tree is created automatically for the results of a free **text** search
279. A method according to claim 278 and wherein said directory tree enables the user to...

...based on several subjects.
285. A method according to claim. 278 and wherein said directory tree is **organized** , and the information items are **sorted** into the directory trec, based on commonality metric that involves several terms.
286. A method according to claim 278 and wherein said directory tree is **organized** , and the information items are **sorted** into the directory tree, based on a commonality metric that involves terms that were not specified by...

...the user's query.
287. A method according to claim 278 and whercin said directory tree is **organized** , and the information items are **sorted** into the directory tree, based on a cornrnonality metric that involves a plurality of terms.
0 288. A method according to claim. 278 and wherein said directory tree is **organized** , and the information item.s are **sorted** into the directory tree, based on a metric of lack of cornmonality between. information items.
72
. A method according to claim 278 and wherein said directory tree is **organized** , and the information items are **sorted** into the directory tree, in an iterative manner where information items are added or removed from clusters...

...of clusters, Create several alternative groups of clusters and select said best group.
298. A system for **organizing** items comprising:
a cluster generator operative to build up clusters of items, each item. having information associated...each other than they have in common with items outside of the cluster. 299. A system for **organizing** items according to claim 298 and wherein said metric is a commonality metric. -)00. A system for **organizing** items according to claim 298 and wherein said metric is a similarity metric. 31 0 1. A system for **organizing** items according to claim 298 and wherein said metric is a non-conunonality metric. 302. A system...

...to claim 298 and wherein said metric is a non-similarity metric. 3 0 '). A system for **organizing** infonnation according to claim 298 and wherein cach item includes at least one descriptor and said metric...

...of claims 304 - 269 and wherein calculating said similarity score includes assigning at least one of a **match** count and an umatch count to a pair of items. 311. A system aceording to claim. 3 1 0 and also comprising weighting at least one of said **match** count and said unrnatch count. -3 12. A system according to claim 310 and wherein said metric includes a metric which is equal to the weighted **match** count.
no
313. A system, according to claim, 3 1 0 and wherein. said metric includes a...most:
80
preferred cluster has the highest Cluster Quality Metric, of all possible first clusters available for **comparison** . 347. A system according claims 337 - 345 and wherein a structure of clusters is presented to the...a set nuraber of iterations have taken place. 362. A system aceording to claim 359 and wherein **limitation** of calculations 1 0 to qualified descriptors are used for calculating a Cluster Quality Metric CQM:
CQM...

...unique descriptors of the items of the collection are identified;
no
Step (b). The identified descriptors are **ranked** according to their
83
popularity in the collection;
Step (c). A "base (inverted exclamation mark)tem" is chosen as a first: item of a "base cluster";
Step (d). A plurality of " **cOmparison** items" are chosen;
Step (e). The base item is considered to be a first item in a "base cluster", and cach **comparison** item. is considered to be a first item. in a " **comparison** cluster";
Step (f). The base cluster, now including all items of the collection havina a hiaher gravity score with respect to the base cluster than with respect to any of the **comparison** clusters, is retained as the desired preferred cluster for the collection. 1 0 -3 6 4. A system. according to claim 359 and wherein said identified unique descriptors are the highly **ranking** descriptors and wherein. descriptors that exist in many items of the collection of items are **ranked** above descriptors existing in few items of the collection. 1 5 365. A system according to claim 364 and wherein. each descriptor receives a **rank** score equal to the number of items of the collection in which that descriptor exists. 366. A system. according to claim 365 and wherein said **ranking** is influenced by a weierhting factor dependent on some characteristics of the descriptors J67. A system according to clairn. 366 and wherein. said **ranking** is influenced by a weighting factor dependent on some characteristics of the items in which they appear. 368. A system. according to claim, 367 and wherein. said **ranking** is influenced by a weighting factor dependent on some characteristics of descriptors of items having few descriptors...

...of items having inany descriptors. 3 0 369. A system according to clairn. 368 and wherein said **ranking** is influenced by a weighting factor dependent on some characteristics of descriptors which are nouns are given said base item is chosen as that item having the highest- **ranking** combination of high- **ranking** descriptors.
A system according to claim 370 and wherein said **ranking** is accomplished by first calculating an item score for each item, which is the sum of the...

...item score.
3 73. A system according to any of claims 298 - 332 and wherein a first **comparison** item is an item having a high item score, yet also having a low similarity
score when **compared** to the base item; and
additional **comparison** items are chosen, being items having a high item

score, yet also having a low similarity score when **compared** to the base item and further having a low similarity score when **compared** to all previously chosen **comparison** items.
j 74. A system according to claim 363 the method comprising:
selecting a base cluster and a plurality of **comparison** clusters, each. of these clusters having a single item, wherein in step (e) a gravity score is...

...for each item of the collection with respect to said base cluster and with respect to cach **comparison** cluster, and each item is added to said cluster with respect to which it has the highest...

...claims 298- 376 and wherein a directory
tree is created automatically for the results of a free **text** search
378. A methold according to claim, 377 and wherein said directory tree enables 0 the user...

...method according to claim 377 and wherein said directory tree is oruanized, and the information items are **sorted** into the directory tree, based on
Cn
commonality metric that involves several terms.
86
. A method according to claim 377 and wherein said directory tree is **organized** , and the information items are **sorted** into the directory tree, based on a coinmonality metric that involves terms that were not specified by...

...the user's query.
-)86. A method according to claim 377 and wherein said directory tree is **organized** , and the information items are **sorted** into the directory tree, based on a commonality metric that involves a plurality of terms. 1 0 387. A method according to claim 377 and wherein said directory tree is **organized** , and the information items are **sorted** into the directory tree, based on a metric of lack of coramonality between information items.
388. A method according to claim. 377 and wherein said directory tree is 1 5 **organized** , and. the information items are **sorted** into the directory tree, in an iterative manner where information items are added or removed from clusters...
...of clusters, create several alternative groups of clusters and select said best group.
397. A system for **organizing** information comprising:
a cluster cracker, breaking down clusters of information items; and
a dynarnic metric evaluator, during...each other than they have in corarnon with items outside of the cluster. 398- A system for **organizing** infonnation according to claim 397 and wherein said metric is a cornrnonality metric. 399- A system for **organizing** inforination according to claim 397 and wherein said metric is a similarity metric. 3 0 400. A system for **organizing** information according to claim 397 and wherein said metric is a non-commonality metric.
88
. A system for **organizing** information aceording to claim 397 and wherein
said metric is a non-similarity metric. 402. A system for **organizing** information according to claim 397 and wherein each (inverted exclamation mark)tein includes at least one descriptor...a most
preferred cluster has the highest Cluster Quality Metric of all possible first, clusters available for **comparison** . 446. A method according claims 436 - 444 and wherein a structure of clusters is presented to the ...a set number of iterations have taken place. 461. A method according to claim 458 and wherein **limitation** of calculations to qualified descriptors are used for calculating a Cluster Quality Metric CQM:
CQM = aY + bY...

...All unique descriptors of the items of the collection are identified;

Step (b). The identified descriptors are **ranked** according to their popularity in the collection;
Step (c). A "base (inverted exclamation mark)tem" is chosen as a first, item of a "base cluster"; Step (d). A plurality of " **comparison** (inverted exclamation mark)tem? are chosen; Step (e). The base item is considered to be' a first item in a "base cluster", and each **comparison** item is considered to be a first item in a " **comparison** cluste?';
Step (f). The base cluster, now including all items of the collection having a higher gravity score with respect to the base cluster than with respect to any of the **comparison** clusters, is retained as the desired preferred cluster for the collection. 4 6 3. A method according to claim 458 and wherein said identified unique descriptors are the highly **ranking** descriptors and wherein descriptors that exist in many items of the collection of items are **ranked** above descriptors existing in few items of the collection. 464. A method according to claim 463 and wherein each descriptor receives a **rank** score equal. to the number of items of the collection in which that descriptor exists. 465. A method according to claim 463 and wherein the **ranking** is influenced by a weighting factor dependent on some characteristics of the descriptors
no
466. A method according to claim 465 and wherein the **ranking** is influenced
98
by a weighting factor dependent on some characteristics of the items in which they appear. 467. A method according to claim 466 and wherein the **ranking** is influenced by a weighting factor dependent on some characteristics of descriptors of items having few descriptors...

...than descriptors of items having many descriptors. 468. A method according to clairn. 467 and wherein. the **ranking** is influenced by a weighting factor dependent on some characteristics of descriptors which are nouns are given...

...according to claim 468 and wherein said base item is chosen as that item having the highest- **ranking** combination of high- **ranking** descriptors. 470. A method according to claim. 469 and wherein said **ranking** is accomplished by first calculating an item score for each item, which is the sum of the...
...the highest item score.
472. A method according any of claims 397 - 471 and wherein a first: **comparison** itern. is an itern. having a high itern. score, yet also having a low similarity
score when **compared** to the base item; and
additional **comparison** items are chosen, being items having a high item score, yet also having a low similarity score when **compared** to the base item and further having a low similarity score when **compared** to all previously chosen **comparison** jO items.
99
. A method according to claim 462 the method comprising:
selecting a base cluster and a plurality of **comparison** clusters, each of
these clusters having a single item, wherein in step (e) a gravity score is...

...for cach item of the collection with respect to said base cluster and with respect to each **comparison** cluster, and each item is added to said cluster with respect to which it has the highest...

...claims 397 - 475 and wherein a directory
tree is created autornatically for the results of a free **text** search
477. A method according to claim 476 and wherein said directory tree enables the user to...

...several subjects.
w
483. A method according to claim 476 and wherein said directory tree is
0 **organized** , and the information items are **sorted** into the directory

tree, based On
t)
conunonality metric that involves several terms.
484. A method according to claim 476 and wherein said directory tree is **organized** , and the information items are **sorted** into the directory tree, based on a
i-7
1 5 cominonality metric that involves terms that...

...the user's query.
485. A method according to claim 476 and wherein said directory tree is **organized** , and the information items are **sorted** into the directory tree, based on a coil-imonality metric that involves a plurality of terms.
486. A method according to claim 476 and wherein said directory tree is **organized** , and the information items are **sorted** into the directory tree, based on a metric of lack of cornmonality between information items.
487. A method according to claim 476 and wherein said directory tree is **organized** , and the information items are **sorted** into the directory tree, in an iterative inanner where information items are added or removed from clusters...according to claira 496 and wherein
tD
said metric is a coramonality metric. 498. A system for **organizing** information according to claim 496 and wherein said metric is a similarity metric.
0
499. A system for **organizing** information according to claim 496 and wherein said metric is a non-commonality metric. 500. A systera for **organizing** information according to claim 496 and wherein 5 said metric is a non-similarity metric. 501. A system for **organizing** information according to claim 496 and wherein each item includes at least one descriptor and said metric...

...of claims 502 - 427 and wherein calculating said similarity score includes assigning at least one of a **match** count and an un- **match** count to a pair of items. 509. A system aecording to claim 508 and also comprising weighting at least one of said **match** count and said unmatch count.
510. A system according to claim 508 and wherein said metric includes a nietric which is equal to the weighted **match** count.
511. A system according to claim 508 and wherein said metric includes a metric which is...a most
preferred cluster has the highest Cluster Quality Metric of all possible first clusters available for **comparison** . 545. A system according claims 535 - 543 and wherein a structure of clusters is presented to the...a set number of iterations have taken place. 560. A system aecording to claim 557 and wherein **limitation** of calculations
111
to qualified descriptors are used for calculating a Cluster Quality Metric CQM:
CQM = aX...

...All unique descriptors of the items of the collection are identified;
Step (b). The identified descriptors are **ranked** according to their popularity in the collection;
Step (c). A "base item` is chosen as a first item of a "base cluster`;
Step (d). A plurality of " **comparison** item? are chosen;
Step (e). The base item is considered to be a first item in a "base cluste?, and cach **comparison** item is considered to be a first item in a " **comparison** cluste?;
Step (f). The base cluster, now including all items of the collection having a higher gravity score with respect to the base cluster than with respect to any Of
ZD
the **comparison** clusters, is retained as the desired preferred cluster

for the collection.
'10
112
. A system. according to claira 557 and wherein said identified unique descriptors are the highly **ranking** descriptors and wherein descriptors that exist in many items of the collection of items are **ranked** above descriptors existing in few items of the collection. 5 6 3. A system according to claim 562 and wherein each descriptor receives a **rank** score equal to the number of items of the collection in which that descriptor exists. 564. A system according to claim 563 and wherein said **ranking** is influenced 1 0 by a weighting factor dependent on some characteristics of the descriptors 565. A system according to claim 564 and wherein said **ranking** is influenced by a weighting factor dependent on some characteristies of the items in which they appear.
1 5
566. A system according to claim 565 and'wherein said **ranking** is influenced by a weighting factor dependent on some characteristics of descriptors of items having few descriptors...

...than descriptors of items having many descriptors. 567. A system. according to claim 566 and wherein said **ranking** is influenced by a weighting factor dependent on some characteristics of descriptors which are nouns are given...

...according to claim 567 and wherein. said base item. is chosen as that item having the highest- **ranking** combination of high- **ranking** descriptors.
569. A system according to claim 568 and wherein said **ranking** is accomplished by first calculating an item, ...the highest item score.
571. A system according any of claims 496 - 570 and wherein a first **comparison** item is an item having a high item score, yet also having a low similarity
score when **compared** to said base item; and
additional **comparison** items are chosen, being items having a high item seore, yet also havina a low similarity score when **compared** to the base item and further having a low similarity score when **compared** to all previously chosen **comparison** items.
572. A system according to claim 561 the method comprising
selecting a base cluster and a plurality of **comparison** clusters, each of these clusters having a single item, wherein in step (e) a gravity score is...

...for each item of the collection with respect to said base cluster and with respect to each **comparison** cluster, and each item is added to said cluster with respect to which (inverted exclamation mark)t...

...496 - 574 and wherein a directory
jO tree is created automatically for the results of a free **text** search
576. A system according to claim 575 and wherein said directory tree enables
114
the user...

...based on several subjects.
582. A system according to claim 575 and wherein said directory tree is **organized** , and the information items are **sorted** into the directory tree, based on conimonality metric that involves several terms.
583. A system according to claim 575 and wherein said directory tree is orcranized, and the information items are **sorted** into the directory tree, based on a
Ib
commonality metric that involves terms that were not specified...

...the user's query.
584. A system according to claim 575 and wherein said directory tree is **organized** , and the information items are **sorted** into the directory tree, based on a commonality metric that involves a plurality of terms.

. A...

...according to claim 575 and wherein said directory tree is or,ganized, and the information items are **sorted** into the directory tree, based on a metric of lack of commonality between infonnation items.

586. A system according to claim 575 and wherein said directory tree is **organized** , and the information items are **sorted** into the directory tree, in an iterative inanner where information items are added or removed from clusters...

12/5,K/38 (Item 21 from file: 349)
DIALOG(R)File 349:PCT FULLTEXT

00913742

IMPROVED MULTISTAGE INTELLIGENT DATABASE SEARCH METHOD

METHODE PERFECTIONNEE DE RECHERCHE A ETAPES MULTIPLES DANS UNE BASE DE DONNEES INTELLIGENTE

Patent Applicant/Assignee:
NETRICS COM INC, 707 State Road, Suite 212, Princeton, NJ 08540, US, US (Residence), US (Nationality), (For all designated states except: US)
Patent Applicant/Inventor:
YIANILOS Peter N, 215 Arreton Road, Princeton, NJ 08540, US, US (Residence), US (Nationality), (Designated only for: US)
KANZELBERGER Kirk G, 900 Davidson Road, Piscataway, NJ 08854, US, US (Residence), US (Nationality), (Designated only for: US)
Legal Representative:
WOODBRIDGE Richard C (et al) (agent), Woodbridge & Associates, P.C., P.O. Box 592, Princeton, NJ 08542, US,

Patent and Priority Information (Country, Number, Date):
Patent: WO 200246965 A1 20020613 (WO 0246965)
Application: WO 2000US33422 20001208 (PCT/WO US0033422)
Priority Application: WO 2000US33422 20001208
Designated States: AE AL AM AT AU AZ BA BB BG BR BY CA CH CN CR CU CZ DE DK DM EE ES FI GB GD GE GH GM HR HU ID IL IN IS JP KE KG KP KR KZ LC LK LR LS LT LU LV MA MD MG MK MN MW MX NO NZ PL PT RO RU SD SE SG SI SK SL TJ TM TR TT TZ UA UG US UZ VN YU ZA ZW
(EP) AT BE CH CY DE DK ES FI FR GB GR IE IT LU MC NL PT SE TR
(OA) BF BJ CF CG CI CM GA GN GW ML MR NE SN TD TG
(AP) GH GM KE LS MW MZ SD SL SZ TZ UG ZW
(EA) AM AZ BY KG KZ MD RU TJ TM
Main International Patent Class: **G06F-017/30**
Publication Language: English
Filing Language: English
Fulltext Availability:
Detailed Description
Claims
Fulltext Word Count: 9222

English Abstract
An improved multistage intelligent database search method includes (1) a prefilter that uses a precomputed index to compute a list of most "promising" records that serves as input to the original multistage search method, resulting in dramatically faster response time; (2) a revised polygraph weighting scheme correcting an erroneous weighting scheme in the original method; (3) a method for providing visualization of character matching strength to users using the bipartite graphs computed by the multistage method; (4) a technique for complementing direct search of textual data with search of a phonetic version of the same data, in such a way that the results can be combined; and (5) several smaller improvements that further refine search quality, deal more effectively with multilingual data and Asian character sets, and make the multistage method a practical and more efficient technique for searching document repositories.

French Abstract
L'invention concerne une methode perfectionnee de recherche a etapes multiples dans une base de donnees intelligente, comprenant (1) un pre-filtre qui utilise un index pretraite pour analyser une liste des enregistrements les plus <= prometteurs >=, cette listeservant de donnee d'entree pour la methode d'origine a etapes multiples, d'ou un temps de reponse sensiblement plus rapide; (2) un systeme de ponderation polygraphique revise, servant a corriger un systeme de ponderation errone dans la methode d'origine; (3) un procede de visualisation du potentiel de correspondance des caracteres pour les utilisateurs appliquant les graphes bipartis traites par la methode a etapes multiples; (4) une

technique pour realiser une recherche directe de donnees textuelles, avec recherche d'une version phonetique des memes donnees, de sorte que les resultats peuvent etre combines; et (5) de nombreux perfectionnements moindres, qui affinent davantage la qualite de la recherche et qui traitent de facon plus efficace les donnees multilingues et les ensembles de caracteres asiatiques, faisant ainsi de cette methode a etapes multiples une technique pratique et plus efficace pour chercher des referentiels documentaires.

Legal Status (Type, Date, Text)
Publication 20020613 A1 With international search report.
Examination 20021010 Request for preliminary examination prior to end of 19th month from priority date

Main International Patent Class: **G06F-017/30**
Fulltext Availability:
Detailed Description

Detailed Description
... close to the beginning of the record (at or close to left-alignment).

After the optimal polygraph **matching** has been determined between a **record** and the query, 5 two small penalty values are added to the total **match** cost of the **record** . One is based on the final query alignment chosen by F3, and is called the "alignment penalty." The other is based on the **record** **length** , and is called the " **record** - **length** penalty." The total penalty added to the **match** cost is small enough to affect the **ranking** only among **records** that have exactly the same similarity with the query (as expressed by the **match** cost). The two penalty values themselves are calibrated in such a way that the alignment penalty takes precedence over the **record** - **length** penalty. In any group of output **records** having exactly the **same** **total** **match** cost, the alignment penalty will cause the **records** to be **sorted** according to query alignment. **Records** having both the same **match** cost and the same query alignment will occur adjacent to each other, and the **record** - **length** penalty will cause this subset of **records** to be **sorted** according to increasing **record** **length** . This generally results in the most natural-seeming **order** of **records** in the output list.

In order to ensure that the total penalty added to the match cost...

12/5,K/55 (Item 38 from file: 349)
DIALOG(R)File 349:PCT FULLTEXT

00813195 **Image available**
CACHING OF OBJECTS IN DISK-BASED DATABASES
MISE EN ANTEMEMOIRE D'OBJETS DANS DES BASES DE DONNEES STOCKEES SUR DISQUE
Patent Applicant/Assignee:
TELEFONAKTIEBOLAGET LM ERICSSON (publ), S-126 25 Stockholm, SE, SE (Residence), SE (Nationality)
Inventor(s):
RONSTROM Ulf Mikael, Ornbacken 14, S-126 51 Hagersten, SE,
Legal Representative:
FUCHSLE Klaus (et al) (agent), Hoffmann . Eitle, Arabellastrasse 4, 81925 Munchen, DE,
Patent and Priority Information (Country, Number, Date):
Patent: WO 200146809 A1 20010628 (WO 0146809)
Application: WO 2000EP12356 20001207 (PCT/WO EP0012356)
Priority Application: DE 19961499 19991220
Designated States: AE AG AL AM AT AU AZ BA BB BG BR BY BZ CA CH CN CR CU CZ DE DK DM DZ EE ES FI GB GD GE GH GM HR HU ID IL IN IS JP KE KG KP KR KZ LC LK LR LS LT LU LV MA MD MG MK MN MW MX MZ NO NZ PL PT RO RU SD SE SG SI SK SL TJ TM TR TT TZ UA UG UZ VN YU ZA ZW
(EP) AT BE CH CY DE DK ES FI FR GB GR IE IT LU MC NL PT SE TR
(OA) BF BJ CF CG CI CM GA GN GW ML MR NE SN TD TG

(AP) GH GM KE LS MW MZ [redacted]SL SZ TZ UG ZW
(EA) AM AZ BY KG KZ MD RU TJ TM
Main International Patent Class: **G06F-012/12**
Publication Language: English
Filing Language: English
Fulltext Availability:
Detailed Description
Claims
Fulltext Word Count: 11143

English Abstract
a data processing device (DPD) comprises a main memory (MM) and a processing means (PM). Data from a data base system (DBS) is stored as pages in the main memory (MM). During processing of the individual objects (OB) of the pages (P) the access frequency to each object (OB) stored in the main memory (MM) is determined. Objects having similar access frequencies are collected in the same data storage section of the main memory (MM). In particular, data objects (OB) can be moved to higher order data storage sections to which a higher access frequency range has been assigned. Thus, data which is more frequently used by the processing means (PM) stays in the main memory (MM) longer and data objects which are not so frequently used are transferred back to the data base or are overwritten earlier. Thus, an efficient usage of the memory space and a reduction of the access time to move frequently used data objects can be achieved.

French Abstract
L'invention concerne un dispositif de traitement de donnees (DPD) comprenant une memoire principale (MM) et un organe de traitement (PM). Les donnees issues d'un systeme de base de donnees (DBS) sont stockees dans la memoire principale (MM) sous forme de pages. Lors du traitement de chacun des objets (OB) desdites pages (P), le dispositif determine la frequence d'acces relative a chaque objet (OB) stocke dans la memoire principale (MM). Les objets assortis de frequences d'acces similaires sont sous forme de pages rassembles dans la meme section de stockage des donnees de la memoire principale (MM). Les objets (OB) de donnees peuvent notamment etre deplaces vers des sections de stockage des donnees d'ordre superieur affectees d'une plage de frequences d'acces superieure. Ainsi, les donnees les plus frequemment utilisees par l'organe de traitement (PM) sejournent plus longtemps dans la memoire principale (MM), tandis que les objets de donnees qui ne sont pas utilises avec la meme frequence sont soit retransferes dans la base de donnees soit ecrases au prealable. Cela permet d'utiliser efficacement l'espace memoire et de reduire le temps d'acces necessaire au deplacement des objets de donnees frequemment utilises.

Legal Status (Type, Date, Text)
Publication 20010628 A1 With international search report.
Publication 20010628 A1 Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.
Examination 20011004 Request for preliminary examination prior to end of 19th month from priority date

Main International Patent Class: **G06F-012/12**
Fulltext Availability:
Claims

Claim
... physical
reference to an object using n index structure; and
Fig. 5b shows an updating procedure using **file** description attributes.
It should be noted that in the drawings the same or similar reference numerals denote...
...data stored on a primary memory device ID, said data in said primary memory device D being **organized** as a

plurality of data blocks each consisting of one or more data objects OB, comprises...

...write means R/W is adapted for writing data objects whose determined access frequency falls in a **predetermined** access frequency **range** in data regions belonging to the ...data sections I may be the same. However, they may be different. only for illustration purposes the **size** of the respective data storage sections PCS-j and RDS-i is shown to be the same or may be different. And only for illustration purposes the **size** of the respective data storage sections PCS-j and RDS-i is shown to be the same...

...predetermined "heat level" to each data storage section.essentially means assigning to each data storage section a **predetermined** access frequency **range** pch-1,...pch j .... pch-J; rdh-1,... rdh-i,... rdh-I (pch: page cache heat; rdh: resident data heat). In the simplest case, as shown in Fig. 2, the **predetermined** read access frequency **range** may only be a single value.
According to the invention each data object OB stored in the...means R/W to write or to collect data objects whose determined access frequency fall in a **predetermined** read access frequency **range** in data regions belonging to the same data storage section.
As shown in Fig. 2, each data...

...j, RDS-i,
comprise a number of data regions PCSP, RDSP. If the "heat" increases with higher **order** data storage sections, as indicated in Fig. 2, it is therefore possible to collect objects OB4, OB5...

...cache memory PCS a hierarchy of overwriting data regions (pages) can be assigned such that the lower **order** data regions, e.g. of data storage section PCS-1, are the first ones which are overwritten...as to whether the combination of two or more objects can be moved up to a higher **order** data storage section.
Furthermore, it is also possible to first load the data, of one page only...
...heat" via the read/write access frequency on a data object level is that data objects of **comparable** **relevance** for the processing means can be collected in the same data storage region PCS-j, RDS-i. Thus, the processing means PM can overwrite data regions of data storage sections having a lower **rank** earlier than data regions of data storage sections having a higher **rank** . Therefore, the main memory MM is not overduely occupied by data (data objects) which are not frequently...it's importance (as indicated by the reduced access frequency) and may actually migrate to a lower **order** data storage section.
If the main memory Mm is provided with the page cache sections as well...

...data memory. As seen in Fig. 3, each data storage region has preferably assigned to it a **predetermined** access frequency **range** pch-1, pch-2; rdh-1, rdh For illustration purposes only two data storage sections are shown. Each access frequency range has an upper and a lower access frequency **threshold** value pch-11ow, pch-1up, pch-21ow, pch-2up; rdh-110w, rdh-1up, rdh-21ow, rdh-2up...

...th data storage section and each access frequency range comprises an upper and an lower access

frequency **threshold** v[redacted]e, wherein said read/write mea[redacted]R/W is adapted to move a data object of the...

...storage section when the access frequency of said data object is greater than said upper access frequency **threshold** value and/or to move a data object of the (i+1)-th data storage section from...

...storage section when the access frequency of said data object is smaller than said lower access frequency **threshold** value.
As shown in Fig. 4, in step ST1 the main memory MM is divided in two...Each data storage section can comprise the same number of data regions PCSP, RDSP (for example the **size** of one page) or can comprise a different number of data regions.
In step ST3 new data...

...frequency value for each object OB is calculated.
In step STS the data object access frequency is **compared** with a respective **threshold** value contained within each access frequency range. The thresholds can for example be the upper or lower...

...specific data object OB there are five possibilities where it could be moved depending on it's **relevance** or access frequency.
Firstly, the object OB* can be moved to the data storage section PCS Furthermore...

...object OI3* in Fig. 3. If it has been determined in step ST5 (on the basis of **comparing** the respective data object access count with the corresponding **threshold**) that the object should be moved, then the respective moving of the object takes place in step...

...in the resident data memory of the main memory MM has data regions each corresponding to a **same** access frequency **range** indicated with pch-2, pch-1; rds-2, rds-1 in Fig. 3. However, in **order** to allow a finer discretesation, it is also possible that the data regions themselves are hierachically **arranged** such that even a movement of a particular data object OB within the same data storage section...is, when an object 03 in the lower data storage section RDS-1 exceeds it's upper **threshold** value rds-1Up it is moved to the next higher data storage region, e.g. RDS However...

...count falls below a value which is lower than rds-1Up. That is, the upper access frequency **threshold** of an i-th data storage section RDS-i can be identical to the lower access frequency **threshold** value of the (i+1)-th data storage region RDS-i or not. Thus, a hysteresis is...to be retrieved a page number (page reference) and a page index is needed (if objects are **organized** as disc structure). Users of a database do not normally provided references to objects in memory but...

...processing means PM an index INX (i.e. a data structure reference table) must be provided in **order** to map this logical reference to a physical reference (page reference and page index) pointing to the...
...11c ID111 may be located. Note that the physical reference p ID1 can equally refer to the **position** in the disc database DB or in the main memory sections, for example in the page cache...

...data object. This becomes even worse, if the data

object is stored back [redacted] at the original **position** p [redacted] in the disc database DB (see ST2) but at the corresponding page p ID* in...find the customer object.
Whilst Fig. 5a shows an example where the index structure INX and the **record** storage is separated, i.e. where the index contains a reference to the **record** and where this reference is updated when the object is moved, Fig. 5b shows a further example, where the index structure and **record** structure is combined, i.e. where the **record** storage is separated into several parts.
As shown in Fig. 5b, the **record** storage contains a first (resident) part with the **file** descriptive attributes (each for example 32 bits) in the **file** descriptive part FDP. The **file** content, i.e. the data object itself (=- 10 kB) is stored in the disc part. One of the **file** descriptive attributes FD1 always points to the respective storage location of the **file** content in the disc part. When there is a movement of the **file** content part, i.e. the object, according to the arrow M, then it is this **file** descriptive part FD1 which is updated with the new location of the **file** content in the disc part. The first part(resident part) can either be referenced by an index...

...as illustrated in Fig. 5b or can be part of the index structure INX (i.e. the **file** descriptive part FDP) is located in the right-hand column p ID of the index structure INX...

...a reference in the key, i.e. in the logical reference c ID1. In all cases the **file** content can be moved as long as references in the first (resident) part are updated.
Thus, according resident part FDP can be stored in the resident data section RDS and the **file** content FT can be stored in the resident data section RDS.
FOURTH EMBODIMENT
In the first to...of computer instructions is also page-based and "code" pages are sent to a so called swap **file** on disc. That is, the primary memory device can be the swap **file** on disc and the second memory device can be a page cache section whilst the data objects...

12/5,K/64 (Item 47 from file: 349)
DIALOG(R)File 349:PCT FULLTEXT

00802534
ANY-TO-ANY COMPONENT COMPUTING SYSTEM
SYSTEME INFORMATIQUE A COMPOSANTS TOUTE CATEGORIE
Patent Applicant/Assignee:
E-BRAIN SOLUTIONS LLC, 1200 Mountain Creek Road, Suite 440, Chattanooga, TN 34705, US, US (Residence), US (Nationality), (For all designated states except: US)
Patent Applicant/Inventor:
WARREN Peter, 1200 Mountain Creek Road, Suite 440, Chattanooga, TN 37405, US, GB (Residence), GB (Nationality), (Designated only for: US)
LOWE Steven, 1625 Starboard Drive, Hixson, TN 37343, US, US (Residence), US (Nationality), (Designated only for: US)
Legal Representative:
MEHRMAN Michael J (agent), Paper Mill Village, Building 23, 600 Village Trace, Suite 300, Marietta, GA 30067, US,
Patent and Priority Information (Country, Number, Date):
Patent: WO 200135216 A2-A3 20010517 (WO 0135216)
Application: WO 2000US31231 20001113 (PCT/WO US0031231)
Priority Application: US 99164884 19991112
Designated States: AE AG AL AM AT AU AZ BA BB BG BR BY BZ CA CH CN CR CU CZ

DE DK DM DZ EE ES FI GB [redacted] GE GH GM HR HU ID IL IN IS JP [redacted] KG KP KR KZ
LC LK LR LS LT LU LV MA MD MG MK MN MW MX MZ NO NZ PL PT RO RU SD SE SG
SI SK SL TJ TM TR TT TZ UA UG US UZ VN YU ZA ZW
(EP) AT BE CH CY DE DK ES FI FR GB GR IE IT LU MC NL PT SE TR
(OA) BF BJ CF CG CI CM GA GN GW ML MR NE SN TD TG
(AP) GH GM KE LS MW MZ SD SL SZ TZ UG ZW
(EA) AM AZ BY KG KZ MD RU TJ TM

Main International Patent Class: **G06F-009/44**
International Patent Class: **G06F-017/22**
Publication Language: English
Filing Language: English
Fulltext Availability:
Detailed Description
Claims
Fulltext Word Count: 275671

English Abstract
A universal data and software structure and method for an Any-to-Any computing machine in which any number of any components can be related to any number of any other components in a manner that is not intrinsically hierarchical and is intrinsically unlimited. The structure and method includes a Concept Hierarchy; each concept or assembly of concepts is uniquely identified and assigned a number in a Numbers Concept Language or uniquely identified in a Non-numbers Concept Language. Each Component or assembly of Components is intrinsically related to all other data items that contain common or related components.

French Abstract
L'invention concerne une structure de donnees et de logiciel universelle ainsi qu'un procede de machine informatique toute categorie dans laquelle des composants, quels qu'ils soient et quel que soit leur nombre, peuvent etre rattaches a d'autres composants, quels qu'ils soient et quel que soit leur nombre, d'une maniere intrinsequement non hierarchisee et intrinsequement illimitee. La structure et le procede comportent une hierarchie conceptuelle; chaque concept ou ensemble de concepts est identifie de maniere unique et recoit un numero dans un langage conceptuel de nombres ou dans un langage conceptuel de non-nombres. Chaque composant ou ensemble de composants est intrinsequement rattache a tous les autres elements de donnees qui contiennent des composants communs ou associes.

Legal Status (Type, Date, Text)
Publication 20010517 A2 Without international search report and to be republished upon receipt of that report.
Search Rpt 20020808 Late publication of international search report
Republication 20020808 A3 With international search report.

Main International Patent Class: **G06F-009/44**
International Patent Class: **G06F-017/22**
Fulltext Availability:
Claims

Claim
... loaded with all knowledge there is on everything, and hence every computer that understands should have the **capacity** to learn. However a' **capacity** to learn' is not something nebulous, but simply a **capacity** to record new word definitions and new rules in memory, and use the new rules, together with...

...it is difficult to find examples of computer Execution Related Memory - at best, such facilities are extremely **limited** . Software does not remember it printed a certain letter last week, but a human does remember that...

...summer time. However, unlike Content Related memory, which is more or less accessible to the user, the **limited** Execution Related Memory facilities that exist, tend to be buried deep inside software and

difficult or difficult. [redacted]f a letter or a book - can sti[redacted]be recorded in the computer in Normal Language - the **same** manner it is recorded today, Under these circumstances, the computer can be ordered in Normal Language to...it

78

possible to identify the types of data needing to be recorded and make arrangements to **record** them. Concept Language Requirements. B. Human Unique Data Specification. Example. The following conversation between a boss and his secretary shows the principal used in identifying a specific **document** uniquely:
Boss to Secretary: Get me the letter.
Secretary to Boss: Which one?
Boss to Secretary: The...

...of 'one' in this instance, and this mismatched with secretaries recording, namely that she has many letters **filed** , not just one. The data mismatch lead to the query ('which ones?'), and when queried, the boss...

...Noticeably, as the Boss added specifications, each added specification was of a new type:
Letter (type of **document**); I (Name of person); Sent to (an Action - Sent); Joe (Name of a person); Bananas (Something inMy' is a word that conveys the concept of 'everything belonging to me.' It is not **limited** in any way. Anything that is the speaker's, is included and falls under the word 'my...

...are given names as in (1), those names are defined using the principle of (2).
4) The **order** in which the co-reducing concepts are stated is of no importance to the meaning conveyed by...

...further described, the one that is to be so described may be coded and indicated by the **order** of the words in the group of Co-Reducing Concepts. It can be easily seen that the **order** is not material to the concept itself
My New York client friend
My friend, a client in...only a fraction of a second. Since someone can only issue a Unique data Specification at a **limited** speed, it is possible to imagine that the item identification process will keep pace with the speed...

...for identifying any stored item in a computer, or any attached item. Essentially, this method creates a **limited** Concept Language that is capable of being used to control a computer in most cases, but will... such an application is attached as Appendix B. However, an application to be programmed can be extremely **limited** , and not require an extensive language review in **order** to identify the possible Concept Hierarchies and hence Data Classes. For example, the application to be programmed might be to enable Normal Language control of a telephone with a **limited** memory. In that case, what the telephone can do is extremely **limited** , and hence the number of possible instructions is **limited** also. Consequently, the Concept Language required is consequently small, and the possible Concept Hierarchies and Data Classes...

...computer:
Step 1) Begin (with the Energy Data Category) and ask 'what can this computer do? In **order** to get answers to Stepl above. Three of the answers might be:
It can print things, store things, fax things. It can make **text** bold
Step 2) Make a list of hardware devices the application is to control. Printerl, Joe's Printer, screen, hard disk etc. Step 3) Assemble the list into groups of **similar** actions. Do this by asking questions of each of the items on the list, to discover if...be located based on a human's Unique Data Specification.

2) They provide the bas[redacted]forenabling general **text** to [redacted]recorded in a computer in such a manner that the computer can be questioned on...

...on the results of its query or question. 1 0 The contribution of Concept Hierarchies to enabling **text** to be questioned as a human
) A Concept Language consisting entirely of numbers - called a Number Concept...

...into an existing Concept Language is
extremely familiar. If the person's understanding of the language is **limited** , or if they believe that some words do not really have specific meanings (some people
believe this...and as being acceptable, some time prior to the date of printing. Hence dictionaries are to some **extent** historical, while a computer should operate in present time. Also it is not the place of a... explained later in the description. An 'Alternate' is defined as 'Two or more
There are only a **limited** number of features that-distinguish in a given piece of text, which meaning of one single spelling...

...in some manner indicate 'telephone' than to words that are names of other things that have the **capacity** of movement. This characteristic is demonstrated in the in following pair of
examples:
'Roam. Yes Roam. The...that it is the relative proximity of the word to a word for something that has the **capacity** to roam that sets the meaning of the word to use. This is type of Context Compression...

...processing can be
0 integrated at any of the following levels:
1) Control only computer. Requires a **limited** Concept Language built from a vocabulary based on the actions that the machine is required to control...of the identification will be no better than the state of the art. This computer can accept **orders** in any machine-readable format such as keyboard, Mouse, Touch-screen, Voice (if Voice Recognition software is...

...by itself, blocks the creation of any
Understanding Computer. This teaching is as follows:
Humans are not **limited** in any way in their **capacity** to think and devise
conce pts, and if a Computer is to Understand, it can not be **limited** either and should 1 0 be able to track and record whatever a human can put into...

...Principle and the application of it enables a computer to follow the human and not impose arbitrary **limits** on him simply to make programmer's life easier. Violating the Unlimited Principle can have small, or...

...in a computer that can not understand.
The Unlimited Principle is stated as:
A computer, within the **limits** of the capacities of its hardware, should never **limit** a user in a manner that he does not **limit** himself.
For example: software that contains a place for three phone number per person only, or only...

...mail addresses, the simple inability to be able to enter only that fourth email address has the **capacity** to half the entirety of the remainder of the computer's understanding. For example, supposing a person...Tables, the appropriate Concept Symbols are placed in the appropriate,
5 corresponding field new Data Relation Table **record** . Hence the new **record** would contain the reference number of **document** X in one of its fields, and an entry
1print' in the Output Data Class Field of the Data Relation table **Record**

corresponding to Output [redacted]ta Class Table, Table 1 above. [redacted]
When the user issues a query for the...revert any action that can be cancelled or reverted; in any case the user expects such a **capacity** , just as he would in the case of a secretary. (That **capacity** is included in the Data RElation Table). The general method of this Any-to-Any machine for...one is in use is detectable not from the word itself, but from characteristics of the surrounding **text** , and again, a characteristic that can be detected and made into a rule. This teaching and understanding...included:
A user says to his computer 'get me the letter I wrote when Joe cancelled his **order** .' The word'When'questions for a value from the Data Category of Time. The query
can be...

...letter I wrote at time Value X
Time Value X,= the Time value for'Joe Cancelled his **order** .'
The two actions (as signified by the words in Italics in the example) are related by a...execution conditions and types of connections between them, as for
example in the order:
'Increase the font **size** of the thing Bill sent me, and fax it to Joe, Take the email about bananas not...

...is improved the enablements of this Any-to-Any machine will enable a computer to understand every **order** , and perform none of them. . In effect, The Data RElation Table extends Concept Language expands the Command...

...This 0 expansion of the definition is partly in the form of an additional Data Relation Table **record** that states the required condition for a module to operate. It has been pointed out that words... shall fire her". When these words are spoken, they can be spoken in a monotone with equal **length** pauses- between words, and still be correctly understood: 'I..am..angry..jili..is..difficult..i.shall..fire ...Category, Data Class Human Life or Data Class Non-human Life, or Matter Data Category with the **capacity** of movement. The Concept Condition Rule Method creates computer - executable statements that software can test to see...

...value should have associated with it a Concept Symbol signifying that the object it represents has the **capacity** to move. 5 Accompanying software logic can now use the Condition Records of the entered a Complete...result in the user adding
Recording Space in a Computer
The only way a computer can really **record** the data in the Space data Category is in
terms of:
. Coordinates. Coordinates can be used to **record** spaces and shapes
- Names of spaces - f o r example, "New York' is name for a space...

...occupies, it is defined as a Coordinate Pattern whose orientation with respect to gravity falls within specified **limits** . The direction of Gravity is a fundamental orienting factory in all verbal expressions of movement. All shapes...been obvious in the state of the art, where there is generally no place or manner to **record** the relationship between movement and location. In the absence of such an **arrangement** , a computer that is **ordered** to move something - for 5 example a box on a screen - does not **record** where the box was, only where it is now. A computer told that'Joe has moved'will...

...and the Space Data Category
The Unlimited Principle, stated previously, is as follows:
A computer, within the **limits** of the capacities of its hardware, should never **limit** a user in a manner that he does not **limit** himself. Most state of the art software violates this principle when dealing with multi-data type called...The 'Any to Any' principle is stated as:
A Human being relates anything to anything within the **limits** of

physical **capacity** , functionality or agreements. Language is itself can be described as any-to-any data transmission system. Any...

...computer does not apply the Any to Any Principle, 5 it will violate the Unlimited Principle by **limiting** the user where he is not himself **limited** . To the degree that it does so, it will not 'understand' him., or he understand it. There...Pinch, but she is not all of Klein and Pinch. if this is not done, then an **order** to 'Contact Klein & Pinch' will fail, because, even thought he computer has a recording for Miss Jones...a number. 2) If a database is used to store data, then it needs to have the **capacity** to rename its fields depending on the data being displayed.
0 Components of an Address - Coordinates in...

...an Address - Handling Words in Addresses. as an Any to Any display problem. In this manner, a **file** in the prior art sense of the word, does not exist. A photograph may bestored in one place, numbers data in another, and **text** elsewhere - each item is stored separately. A' **Document** 'or a'Lefter'then, simply becomes an output-time display assembly job, of assembling the right components...

...correct spatial relationships to one another, and this, when seen on the screen or printed, is 'the **document** ' or 'the letter'. Thus Any output can consist of Any combination of Any data type. Thus Concept...the rule, the correct Concept Language value is provided to the Execution processing software. Obviously, only imagination **limits** the number of different ways in which different, unique values can be assigned to different meanings of...

...meanings could equally well be represented by an electronic signal or an audio signal, or a different **length** or format of light pulse, or a light pulse and a radar pulse. Equally, a Concept Language...machine adds functionality to all existing software, and adds usefulness to all existing data. The user is **limited** by the available processing power, by available storage **capacity** and by the physical **limits** of peripherals that are connected to, or in some manner accessible to the Any-to-Any machine, and by installed logics, but is not otherwise **limited** . As will be seen, when the Any-to-Any machine is used to 5 create a data...between, and create its own output in any installed language in any installed language domain. The Term ' **Order** Execution Processing' is introduced and defined as: 'the collection of mechanisms needed 'to take output from the...

...Visual Interface input or both, and use them execute and control the Execution of the user's **orders** . The definition is continued to include 'those mechanisms that supply output to the Grammar Formatting system of ...software Component.' However, as far as a programmer is concerned, a block of code that may make **text** bold, may in fact be composed of several different pieces of code, each of which performs one of the several steps required to make **text** bold. One block of code, for example, may do one action such as'get name of the...

...value for font name in use' and another'place value in buffer A' etc. Thus, the 'make **text** bold'code can be broken down into three constituent Component blocks of code. Supposing one of these...

```
Set     Items   Descripti
S1    2842320   DOCUMENT OR RECORD? ? OR FILE? OR TEXT?
S2    8223531   SORT? OR ORDER? OR RANK? OR ARRANG? OR ORGANIZ? OR RELEVAN-
            CE? OR POSITION?
S3    8705474   SIZE? OR SIZING OR THRESHOLD OR LENGTH OR CAPACITY? OR LIM-
            IT? OR EXTENT
S4    8495914   COMPAR? OR MATCH? OR EQUATE?
S5    9034004   SIMILAR OR ALIKE OR COMPARABLE OR CORRESPOND? OR EQUIVALENT
             OR SAME OR EQUAL OR IDENTICAL OR PREDETERMIN? OR AROUND
S6    8639302   RANGE OR EXTENT OR REALM OR SCOPE OR SPHERE OR AMOUNT OR T-
            OTAL OR TOTALITY OR SMALLER OR LARGER
S7      53141   S1 (S) S2 (S) S3
S8       1556   S4 (S) S1 (S) (S5 (3N) S6)
S9        157   S7 (S) S8
S10       148   S9 NOT PY>2002
S11       131   S10 NOT PD>20020122
S12        96   RD (unique items)
File   2:INSPEC 1969-2004/Jun W4
         (c) 2004 Institution of Electrical Engineers
File   6:NTIS 1964-2004/Jul W1
         (c) 2004 NTIS, Intl Cpyrght All Rights Res
File   8:Ei Compendex(R) 1970-2004/Jun W4
         (c) 2004 Elsevier Eng.  Info. Inc.
File  34:SciSearch(R) Cited Ref Sci 1990-2004/Jul W1
         (c) 2004 Inst for Sci Info
File  35:Dissertation Abs Online 1861-2004/May
         (c) 2004 ProQuest Info&Learning
File  65:Inside Conferences 1993-2004/Jul W1
         (c) 2004 BLDSC all rts. reserv.
File  92:IHS Intl.Stds.& Specs. 1999/Nov
         (c) 1999 Information Handling Services
File  94:JICST-EPlus 1985-2004/Jun W2
         (c)2004 Japan Science and Tech Corp(JST)
File  95:TEME-Technology & Management 1989-2004/Jun W1
         (c) 2004 FIZ TECHNIK
File  99:Wilson Appl. Sci & Tech Abs 1983-2004/Jun
         (c) 2004 The HW Wilson Co.
File 103:Energy SciTec 1974-2004/Jun B2
         (c) 2004 Contains copyrighted material
File 144:Pascal 1973-2004/Jun W4
         (c) 2004 INIST/CNRS
File 202:Info. Sci. & Tech. Abs. 1966-2004/May 14
         (c) 2004 EBSCO Publishing
File 233:Internet & Personal Comp. Abs. 1981-2003/Sep
         (c) 2003 EBSCO Pub.
File 239:Mathsci 1940-2004/Aug
         (c) 2004 American Mathematical Society
File 275:Gale Group Computer DB(TM) 1983-2004/Jul 07
         (c) 2004 The Gale Group
File 434:SciSearch(R) Cited Ref Sci 1974-1989/Dec
         (c) 1998 Inst for Sci Info
File 647:CMP  Computer Fulltext 1988-2004/Jun W4
         (c) 2004 CMP Media, LLC
File 674:Computer News Fulltext 1989-2004/Jun W3
         (c) 2004 IDG Communications
File 696:DIALOG Telecom. Newsletters 1995-2004/Jul 06
         (c) 2004 The Dialog Corp.
```

12/5/13 (Item 13 from file: 2)
DIALOG(R)File 2:INSPEC

03222455 INSPEC Abstract Number: C88058344
Title: Sorting large files on a backend multiprocessor
Author(s): Beck, M.; Bitton, D.; Wilkinson, W.K.
Author Affiliation: Dept. of Comput. Sci., Cornell Univ., Ithaca, NY, USA

Journal: IEEE Transactions on Computers vol.37, no.7 p.769-78
Publication Date: July 1988 Country of Publication: USA
CODEN: ITCOB4 ISSN: 0018-9340
U.S. Copyright Clearance Center Code: 0018-9340/88/0700-0769$01.00
Language: English Document Type: Journal Paper (JP)
Treatment: Practical (P)
Abstract: The authors investigate the feasibility and efficiency of a parallel **sort** -merge algorithm by considering its implementation of the JASMIN prototype, a backend multiprocessor built around a fast packet bus. They describe the design and implementation of a parallel **sort** utility and present and analyze the results of measurements **corresponding** to a **range** of **file sizes** and processor configurations. The results show that using current, off-the-shelf technology coupled with a streamlined distributed operating system, three- and five-microprocessor configurations, provide a very cost-effective **sort** of large **files** . The three-processor configuration **sorts** a 100-Mb **file** in 1 hr which **compares** well to commercial **sort** packages available on high-performance mainframes. In additional experiments, the authors investigate a model to tune their **sort** software and scale their results to higher processor and network capabilities. (17 Refs)
Subfile: C
Descriptors: database management systems; multiprocessing systems; sorting
Identifiers: backend multiprocessor; parallel sort-merge algorithm; JASMIN prototype; fast packet bus; streamlined distributed operating system
Class Codes: C5440 (Multiprocessor systems and techniques); C6130 (Data handling techniques); C6160 (Database management systems (DBMS))

12/5/14 (Item 14 from file: 2)
DIALOG(R)File 2:INSPEC

02854110 INSPEC Abstract Number: C87022043
Title: Design and evaluation of a parallel sort utility
Author(s): Beck, M.; Bitton, D.; Wilkinson, W.K.
Author Affiliation: Dept. of Comput. Sci., Cornell Univ., Ithaca, NY, USA

Conference Title: Proceedings of the 1986 International Conference on Parallel Processing (Cat. No.86CH2355-6) p.934-41
Editor(s): Hwang, K.; Jacobs, S.M.; Swartzlander, E.E.
Publisher: IEEE Comput. Soc. Press, Washington, DC, USA
Publication Date: 1986 Country of Publication: USA xviii+1051 pp.
ISBN: 0 8186 0724 6
U.S. Copyright Clearance Center Code: 0190-3918/86/0000-0934$01.00
Conference Sponsor: IEEE; Pennsylvania State Univ.; ACM
Conference Date: 19-22 Aug. 1986 Conference Location: St. Charles, IL, USA
Language: English Document Type: Conference Paper (PA)
Treatment: Practical (P)
Abstract: A fundamental measure of processing power in a database management system is the performance of the **sort** utility it provides. When **sorting** a large data **file** on a serial computer, performance is **limited** by factors involving processor speed, memory **capacity** , and I/O bandwidth. An investigation is made of the feasibility and efficiency of a parallel **sort** -merge algorithm through implementation on the JASMIN prototype, a backend multiprocessor built around a fast packet bus. The design and implementation of a parallel **sort** utility are described. The

results of measurements **corresponding** to a **range** of **file sizes** and processor configurations are analyzed, showing that using current, off-the-shelf technology coupled with a streamlined distributed operating system, three- and five-microprocessor configurations provide a very cost-effective **sort** of large **files** . The three-processor configuration **sorts** a 100-Mb **file** in one hour, which **compares** well with commercial **sort** packages available on high-performance mainframes. (16 Refs)

Subfile: C

Descriptors: database management systems; distributed processing; operating systems (computers); parallel algorithms; performance evaluation; sorting

Identifiers: parallel sort utility; processing power; database management system; performance; large data file; processor speed; memory capacity; I/O bandwidth; parallel sort-merge algorithm; JASMIN prototype; backend multiprocessor; fast packet bus; design; implementation; streamlined distributed operating system; 100-Mb file

Class Codes: C4240 (Programming and algorithm theory); C6150J (Operating systems); C6160 (Database management systems (DBMS))

12/5/38 (Item 1 from file: 35)

01917603 ORDER NO: AADAA-I3072155

Sharing and caching characteristics of Internet content

Author: Wolman, Alastair

Degree: Ph.D.
Year: 2002
Corporate Source/Institution: University of Washington (0250)
Chair: Henry M. Levy
Source: VOLUME 63/11-B OF DISSERTATION ABSTRACTS INTERNATIONAL. PAGE 5344. 148 PAGES
Descriptors: COMPUTER SCIENCE
Descriptor Codes: 0984
ISBN: 0-493-92292-X

To improve the performance of Internet content delivery, many techniques exploit sharing: repeated requests to the same object by multiple clients. One widely deployed technique is Web proxy caching, where requests to shared objects are served from a proxy cache instead of the origin server. In this dissertation, we present a network tracing system that enables the study of application-level Internet workloads, and we present three Internet caching studies performed using workloads collected by the tracing system.

The first study investigates Web **document** sharing patterns from an <italic> **organizational** </italic> point of view. We explore the **extent** of **document** sharing both within and across **organizations** . We find that when clients are members of the **same organization** , the **amount** of sharing increases measurably when **compared** with clients that are members of different **organizations** . However, this increase is not large enough to have a significant impact on cache performance.

The second study explores the performance of cooperative Web proxy caching, focusing on the effectiveness of cooperation over a wide range of client population sizes. Allowing proxy caches to cooperate effectively combines the client populations served by those proxies. This provides new opportunities for sharing, and therefore offers the potential to increase cache hit rates. Overall, we find that proxy cooperation provides significant performance benefits only within limited population bounds.

The final study is motivated by the increasing availability of multimedia Internet content, such as streaming audio and video. We compare the workload characteristics of streaming-media content to traditional Web content, and we evaluate the effectiveness of proxy caching and multicast delivery for streaming-media content. We find that these multimedia workloads exhibit strong temporal locality, and we quantify the benefit it provides for caching and multicast delivery.

Finally, we present the design and implementation of our trace

collection system. It uses passive network monitoring to observe all Web traffic generated by the University of Washington client population. Our system employs anonymization safeguards to protect users' privacy. It has been deployed at the University network border for three years, and has scaled to handle a factor of three load increase during that period.

```
Set     Items   Description
S1    1169075   DOCUMENT OR RECORD? ? OR FILE? OR TEXT?
S2    4636328   SORT? OR ORDER? OR RANK? OR ARRANG? OR ORGANIZ? OR RELEVAN-
            CE? OR POSITION?
S3    2621192   SIZE? OR SIZING OR THRESHOLD OR LENGTH OR CAPACITY? OR LIM-
            IT? OR EXTENT
S4    1143479   COMPAR? OR MATCH? OR EQUATE?
S5    4370428   SIMILAR OR ALIKE OR COMPARABLE OR CORRESPOND? OR EQUIVALENT
             OR SAME OR EQUAL OR IDENTICAL OR PREDETERMIN? OR AROUND
S6    2492781   RANGE OR EXTENT OR REALM OR SCOPE OR SPHERE OR AMOUNT OR T-
            OTAL OR TOTALITY OR SMALLER OR LARGER
S7      40460   S1 AND S2 AND S3
S8       1083   S4 AND S1 AND (S5 (3N) S6)
S9         80   S7 AND S8
S10        24   S9 AND IC=G06F?
File 347:JAPIO Nov 1976-2004/Mar(Updated 040708)
         (c) 2004 JPO & JAPIO
File 350:Derwent WPIX 1963-2004/UD,UM &UP=200443
         (c) 2004  Thomson Derwent
```

10/5/5 (Item 5 from file: 347)
DIALOG(R)File 347:JAPIO

02249522 **Image available**
SORTING METHOD FOR VARIABLE **LENGTH** **RECORD**

PUB. NO.: 62-166422 [JP 62166422 A]
PUBLISHED: July 22, 1987 (19870722)
INVENTOR(s): KATO TAKESHI
APPLICANT(s): ALPS ELECTRIC CO LTD [001009] (A Japanese Company or Corporation), JP (Japan)
APPL. NO.: 61-008548 [JP 868548]
FILED: January 18, 1986 (19860118)
INTL CLASS: [4] **G06F-007/24**
JAPIO CLASS: 45.1 (INFORMATION PROCESSING -- Arithmetic Sequence Units); 45.2 (INFORMATION PROCESSING -- Memory Units)
JAPIO KEYWORD: R139 (INFORMATION PROCESSING -- Word Processors)
JOURNAL: Section: P, Section No. 653, Vol. 12, No. 6, Pg. 123, January 09, 1988 (19880109)

ABSTRACT

PURPOSE: To minimize a work area by **comparing** data which are inputted successively on the basis of the predetermined **size** of an array and setting smaller data in the prescribed work area.

CONSTITUTION: The value of input data is set in a **comparison** area A. The 1st data, however, is set in a **comparison** area B. The data in the areas B and A are **compared** with each other and when it is confirmed that the data in the area B is **smaller** or **equal** , the value in the area B is set in a buffer 0 and the value in the area A is moved to the area B. When the value of the area A is smaller than the value in the area B, it is confirmed that the (n)th byte of the word data in the area A is smaller than the word data in the area B, and the word is set in a buffer (n). Said operation is repeated up to the final data and the data are all **sorted** in buffers 0-(n) in predetermined array **order** .

10/5/13 (Item 8 from file: 350)
DIALOG(R)File 350:Derwent WPIX

011815818 **Image available**
WPI Acc No: 1998-232728/199821
Related WPI Acc No: 1999-570571; 1999-609600; 2000-095406; 2004-024612
XRPX Acc No: N98-184363
Record **storing method for storing geographical data on storage medium - by separating geographic data into parcels having desired fill value and dividing** arrangement **that enables their addressing and identification**
Patent Assignee: NAVIGATION TECHNOLOGIES CORP (NAVI-N); ASHBY R A (ASHB-I); ISRANI V S (ISRA-I); NYCZAK G M (NYCZ-I); SMITH N E (SMIT-I)
Inventor: ASHBY R A; BOUZIDE P M; CRANE A I; FERNEKES R P; ISRANI V; JASPER J C; LAMPERT D S; MEEK J A; NYCZAK G M; SMITH N E; ISRANI V S
Number of Countries: 021 Number of Patents: 007
Patent Family:

Patent No	Kind	Date	Applicat No	Kind	Date	Week	
EP 838663	A2	19980429	EP 97308527	A	19971024	199821	B
CA 2219043	A	19980425	CA 2219043	A	19971024	199836	
JP 10312153	A	19981124	JP 97332262	A	19971027	199906	
US 6308177	B1	20011023	US 96740295	A	19961025	200165	
			US 99362947	A	19990728		
CA 2219043	C	20030218	CA 2219043	A	19971024	200327	
EP 838663	B1	20031022	EP 97308527	A	19971024	200373	
			EP 200377520	A	19971024		
DE 69725677	E	20031127	DE 625677	A	19971024	200403	
			EP 97308527	A	19971024		

Priority Applications (No Type Date): US 96740295 A 19961025; US 99362947 A 19990728
Patent Details:

Patent No	Kind	Lan	Pg	Main IPC	Filing Notes
EP 838663	A2	E	58	G01C-021/20	
Designated States (Regional): AT BE CH DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE					
CA 2219043	A			G06F-017/00	
JP 10312153	A		169	G09B-029/10	
US 6308177	B1			G06F-017/30	Div ex application US 96740295 Div ex patent US 5968109
CA 2219043	C	E		G06F-017/00	
EP 838663	B1	E		G01C-021/20	Related to application EP 200377520
Designated States (Regional): AT BE CH DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE					
DE 69725677	E			G01C-021/20	Based on patent EP 838663

Abstract (Basic): EP 838663 A

The method includes separating the number of **records** into first and second groups of **records** so that the **records** in the first group represent physical features having geographic locations encompassed within a first sub-rectangular area and the **records** in the second group represent physical features having geographic locations encompassed within a second sub-rectangular area.

The two sub-rectangular areas are formed by a division at a **position** of a rectangular area that encompasses the locations of the physical features represented by the number of **records** in the first and second groups. The **position** of the division is determined by evaluating a number of trial divisions of the rectangular area, and selecting one of the trial divisions based upon resultant **sizes** of the groups.

The resultant **sizes** of the first and second groupings derived from the evaluation of the trial divisions are **compared** to a first range of **sizes**, and the **records** are into first and second groups based upon at least one of the groups **corresponding** to the first **range** of **sizes**.

ADVANTAGE - Provides potential for enhancing speed and operation of navigation application functions that use geographic data on storage medium. Can up-date real-time traffic information via wireless communication to supplement database installed in vehicle.

Dwg.3/11

Title Terms: **RECORD** ; STORAGE; METHOD; STORAGE; GEOGRAPHICAL; DATA; STORAGE; MEDIUM; SEPARATE; GEOGRAPHICAL; DATA; PARCEL; FILL; VALUE; DIVIDE; **ARRANGE** ; ENABLE; ADDRESS; IDENTIFY
Derwent Class: P85; S02; T01; W06
International Patent Class (Main): G01C-021/20; **G06F-017/00** ; **G06F-017/30** ; G09B-029/10
International Patent Class (Additional): G01C-021/00; **G06F-017/50** ; G06T-001/00; G08G-001/0968; G08G-001/0969
File Segment: EPI; EngPI

10/5/14 (Item 9 from file: 350)
DIALOG(R)File 350:Derwent WPIX

010541997 **Image available**
WPI Acc No: 1996-038951/199604
XRPX Acc No: N96-032839

Sorting **collection of** records **contg. different amounts of data - normalising amount of data in each** record **to value selected from progression of integer powers of two or Fibonacci numbers, partitioning into subsets of** records **each contg.** same amount **of data**, sorting **and merging**

Patent Assignee: DIGITAL EQUIP CORP (DIGI)
Inventor: LAWSON J R

Number of Countries: 001 Number of Patents: 001
Patent Family:

Patent No	Kind	Date	Applicat No	Kind	Date	Week	
US 5440736	A	19950808	US 93157864	A	19931124	199604	B

Priority Applications (No Type Date): US 93157864 A 19931124
Patent Details:

Patent No	Kind	Lan	Pg	Main IPC	Filing Notes
US 5440736	A		7	G06F-017/30	

Abstract (Basic): US 5440736 A

A collection of **records** is stored in RAM and/or a hard disk depending upon storage costs or processing time. The amount of data in each **record** is measured by the number of bytes, unpredictably distributed over a large range of values. The number of bytes in each **record** is normalised to a value chosen from a designated set of values, such as a progression computed as a power of two, a Poisson distribution or Fibonacci series, to a value which is at least as large as the number of bytes in the **record** . E.g. if a **record** contains 11 bytes of data, the normalised **record** has 16 bytes. A **record** retains the original amount of data if the corresp. unnormalised **record** contains 22 bytes. The bytes used to normalise the **records** can be appended to the **record** as null bytes.

The normalised collection can be partitioned into subsets of **records** and **sorted** in parallel to gain time efficiencies.

After **sorting** , the **records** are merged to reconstruct the entire collection into a desired sequence.

ADVANTAGE - Increase space and time efficiencies. Normalised **size** of each **record** is rapidly determined by simple and fast-to-execute bit operations. Incremental difference between successive selected **sizes** is small for smaller values to improve fit of collection where number of small **records** is large, and number of large **records** is small. Storage space in many computers is allocated in quantities expressed as integer powers of two. **Sorting** collection of 16000 **records** , ranging in **size** from 32 to 45000 bytes can be **sorted** in 5 seconds, **compared** with 20 minutes.

Dwg.5/5

Title Terms: **SORT** ; COLLECT; **RECORD** ; CONTAIN; AMOUNT; DATA; NORMALISE; AMOUNT; DATA; **RECORD** ; VALUE; SELECT; PROGRESS; INTEGER; POWER; TWO; FIBONACCI; NUMBER; PARTITION; SUBSET; **RECORD** ; CONTAIN; AMOUNT; DATA; **SORT** ; MERGE
Derwent Class: T01
International Patent Class (Main): **G06F-017/30**
File Segment: EPI

10/5/18 (Item 13 from file: 350)
DIALOG(R)File 350:Derwent WPIX

007500395 **Image available**
WPI Acc No: 1988-134328/198820
XRPX Acc No: N88-102231

Multi-directional scan and print type character generator for printer - produces serial binary stream to print or display in any of 8 combination and progression

Patent Assignee: INT BUSINESS MACHINES CORP (IBMC); IBM CORP (IBMC)
Inventor: FINLAY D E; HANNA S D; STEVENSON D C; VARGA J T
Number of Countries: 004 Number of Patents: 004
Patent Family:

Patent No	Kind	Date	Applicat No	Kind	Date	Week	
EP 267418	A	19880518	EP 87114432	A	19871002	198820	B
US 4841453	A	19890620	US 86929036	A	19861110	198931	
EP 267418	B1	19930818	EP 87114432	A	19871002	199333	
DE 3787073	G	19930923	DE 3787073	A	19871002	199339	
			EP 87114432	A	19871002		

Priority Applications (No Type Date): US 86929036 A 19861[illegible]
Cited Patents: 2.Jnl.Ref; A3...8948; EP 196656; EP 89848; JP 58078279; JP 58121486; No-SR.Pub; US 4000486; US 4079458
Patent Details:

Patent No	Kind	Lan	Pg	Main IPC	Filing Notes
EP 267418	A	E	4		
Designated States (Regional): DE FR GB					
US 4841453	A		36		
EP 267418	B1	E	46	G06K-015/02	
Designated States (Regional): DE FR GB					
DE 3787073	G			G06K-015/02	Based on patent EP 267418

Abstract (Basic): EP 267418 A

The imaging device has a page buffer (160) containing a representation of symbols in a predetermined **order** , each representation including an a/e pointer. A column **position** escape memory (170) specifies for each of the symbol rows, a page buffer pointer to a first symbol in the row and a height, factor identifying a space alloted for representing a symbol row. An address escape memory (150) is provided, having an entry for each different symbol containing a font pointer and a representation of a space alloted for the symbol in orthogonal directions.

A font memory (140) provides a graphic representation of the symbol in at least 2 orientations. A pint commond can be stored which specifies a relaton between symbol rows, scan direction and scan pogresson. An addressing device successively drives diffegivenent font addresses for a symbol until a quantity of data extracted from the font memory beams a specified relation to data extracted from the address escape memory.

ADVANTAGE - Can be used with any paper regardless of printers **limitations** on paper feed ind paper orientation.

Title Terms: MULTI; DIRECTION; SCAN; PRINT; TYPE; CHARACTER; GENERATOR; PRINT; PRODUCE; SERIAL; BINARY; STREAM; PRINT; DISPLAY; COMBINATION; PROGRESS
Index Terms/Additional Words: MULTI; DIRECTION; SCA
Derwent Class: S06; T04
International Patent Class (Main): G06K-015/02
International Patent Class (Additional): **G06F-015/40** ; G06K-015/12
File Segment: EPI

10/5/21 (Item 16 from file: 350)
DIALOG(R)File 350:Derwent WPIX

004622832
WPI Acc No: 1986-126175/198620
XRPX Acc No: N86-093272

Digital data reproduction system - incorporates intermediate data storage to allow time base correction

Patent Assignee: HITACHI DENSHI KK (HITN); HITACHI LTD (HITA)
Inventor: ETO Y; KANADA H; ROKUDA M; UMEMOTO M
Number of Countries: 004 Number of Patents: 006
Patent Family:

Patent No	Kind	Date	Applicat No	Kind	Date	Week	
DE 3539182	A	19860507	DE 3539182	A	19851105	198620	B
GB 2168183	A	19860611	GB 8526814	A	19851031	198624	
JP 61113166	A	19860531	JP 84232503	A	19841106	198628	
US 4700240	A	19871013	US 85795264	A	19851105	198743	
GB 2168183	B	19880901				198835	
DE 3539182	C	19890622				198925	

Priority Applications (No Type Date): JP 84232503 A 19841106
Patent Details:

Patent No	Kind	Lan	Pg	Main IPC	Filing Notes
DE 3539182	A		28		

Abstract (Basic): DE 3539102 C

The system has a number of read-out devices for sequential read-out of individual data sets from the recording medium, fed to a store with a corresp. storage **capacity** . The data sets are subsequently read-out from the store in a preset **order** determined by a read-out address generator.

The input address generator for the store responds to number data obtained from each data set read from the recording medium, to select the write-in addresses.

USE - For time base correction during reproduction of magnetically recorded data. (28pp Dwg.No.0/11

Title Terms: DIGITAL; DATA; REPRODUCE; SYSTEM; INCORPORATE; INTERMEDIATE; DATA; STORAGE; ALLOW; TIME; BASE; CORRECT

Index Terms/Additional Words: VIDEO

Derwent Class: T01; T03; W04

International Patent Class (Additional): **G06F-009/32** ; **G06F-012/02** ; G11B-005/09; G11B-020/12; G11B-027/30; G11C-008/00; H04L-007/00; H04N-005/93

File Segment: EPI

10/5/23 (Item 18 from file: 350)

DIALOG(R)File 350:Derwent WPIX

003107984

WPI Acc No: 1981-L8032D/198146

Data comparator **for rapid retrieval of** records - compares **profile string of patterns with source string of bit patterns to determine** extent **of** matching

Patent Assignee: XEROX CORP (XERO)

Inventor: JONES P F

Number of Countries: 001 Number of Patents: 001

Patent Family:

Patent No	Kind	Date	Applicat No	Kind	Date	Week	
GB 1602591	A	19811111				198146	B

Priority Applications (No Type Date): GB 7824898 A 19780531

Patent Details:

Patent No	Kind	Lan	Pg	Main IPC	Filing Notes
GB 1602591	A		8		

Abstract (Basic): GB 1602591 A

A first addressable store contains a profile string. It provides, in each of a succession of time periods, an output representing the presence and the **position** in, or absence from, the profile string of each bit pattern of the source string. The store is addressable by each bit pattern of the latter combined with the previously identified **position** .

A second addressable store contains data representing the **predetermined extent** of the **matching** . It is addressable by the output of the first store. The output of the second store is used to provide an indication when **matching** has occurred. The second store can also be addressed by the output of a **position** register and a non-**match** counter: both responsive to the output of the first store.

1

Title Terms: DATA; **COMPARATOR** ; RAPID; RETRIEVAL; **RECORD** ; **COMPARE** ; PROFILE; STRING; PATTERN; SOURCE; STRING; BIT; PATTERN; DETERMINE; **EXTENT** ; **MATCH**

Derwent Class: T01

International Patent Class (Additional): **G06F-007/02**

File Segment: EPI